

University of Groningen

Reasoning about local properties in modal logic

Kooi, Barteld; Ditmarsch, Hans van; Hoek, Wiebe van der

Published in:

Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2011

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kooi, B., Ditmarsch, H. V., & Hoek, W. V. D. (2011). Reasoning about local properties in modal logic. In K. Turner, P. Yolum, L. Sonenberg, & P. Stone (Eds.), *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems* (pp. 711-718). IFAAMAS.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Reasoning about local properties in modal logic

Hans van Ditmarsch
Department of Logic
University of Seville
Spain
hvd@us.es

Wiebe van der Hoek
Department of Computer
Science
University of Liverpool, UK
wiebe@csc.liv.ac.uk

Barteld Kooi
Faculty of Philosophy
University of Groningen
the Netherlands
B.P.Kooi@rug.nl

ABSTRACT

In modal logic, when adding a syntactic property to an axiomatisation, this property will semantically become true in all models, in all situations, under all circumstances. For instance, adding a property like $K_ap \rightarrow K_bp$ (agent b knows at least what agent a knows) to an axiomatisation of some epistemic logic has as an effect that such a property becomes *globally* true, i.e., it will hold in all states, at all time points (in a temporal setting), after every action (in a dynamic setting) and after any communication (in an update setting), and every agent will know that it holds, it will even be common knowledge. We propose a way to express that a property like the above only needs to hold *locally*: it may hold in the actual state, but not in all states, and not all agents may know that it holds. We can achieve this by adding relational atoms to the language that represent (implicitly) quantification over all formulas, as in $\forall p(K_ap \rightarrow K_bp)$. We show how this can be done for a rich class of modal logics and a variety of syntactic properties.

Categories and Subject Descriptors

I.2.4 [Artificial Intelligence]: Knowledge Representation Formalisms and Methods—*Modal Logic*

General Terms

Theory

Keywords

Modal logic, Correspondence theory, Canonicity, Local properties, Epistemic logic

1. INTRODUCTION

Modal logic has become *the* framework for formalising areas in computer science and artificial intelligence as diverse as distributed computing [10], reasoning about programs [11], verifying temporal properties of systems, game theoretic reasoning [18], and specifying and verifying multi-agent systems [21]. Regarding the latter example alone, since Moore's pioneering work [14] on knowledge and action, agent theories like intention logic [4] and BDI [15] use

Cite as: Reasoning about local properties in modal logic, van Ditmarsch, van der Hoek, and Kooi, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 711–718.

Copyright © 2011, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

modal logic (where the modalities represent time, action, informational attitudes like knowledge or belief, or motivational attitudes like desires or intentions) to analyse interactions between modalities, like *perfect recall*, *no-learning*, *realism*, or different notions of *commitment*. As for epistemic modal logic, since the seminal work of Hintikka [12], modal epistemic logic has played a key role in knowledge representation, witnessed by its key role for reasoning about knowledge in computer science [7], and artificial intelligence [13]. The current activities in dynamic epistemic logic [1, 19] can be seen as providing a modal logical analysis in the area of belief revision, thereby providing it with a natural basis to do multi-agent belief revision, give an account of the change of higher order information, capture this all in one and the same object language: a modal language, indeed.

The popularity of modal logic in those areas is partly explained by its appealing *semantics*: the notion of state is a very powerful one when it comes to modeling computations of a machine, or describing possibilities that an agent thinks/desires/fears to be possible. Another strong feature of modal logic is its *flexibility*: the fact that temporal, dynamic, informational and motivational attitudes can be represented by modalities does not mean that they all satisfy the same laws. Rather, depending on the interpretation one has in mind, one can decide to either embrace or abandon certain principles for each of the modalities used. Syntactically, this means one assumes a number of axioms or inference rules for a modality or for the interaction of some modalities, and more often than not, this semantically corresponds to assuming some specific properties of the associated accessibility relations.

In the context of epistemic logic for instance, adding specific modal axioms allows one to specify that the knowing agent is *veridical* ($K_ap \rightarrow p$): if agent a knows that p , then p must be true), or that he is positively ($K_ap \rightarrow K_aK_ap$) or negatively ($\neg K_ap \rightarrow K_a\neg K_ap$) *introspective*. Those axioms happen to correspond (in a precise way: correspondence theory for modal logic is already some decades old, cf. [17]) to reflexivity, transitivity and euclidicity of the associated accessibility relation R_a , respectively. Moreover, the axioms are *canonical* for it: adding the syntactic axiom to a modal logic enforces the canonical model for the logic to have the corresponding property, which then in turn implies that completeness of the logic with respect to the class of models satisfying that relational property is guaranteed. At this point, is important to note the difference between $K_ap \rightarrow p$ as a *formula* and that as a *scheme*, or *axiom*: as a formula, it merely expresses that regarding the atom

p , agent a cannot know it without it being true. However, when we assume it as an axiom, or as a scheme, it means that we assume it to hold for every substitution instance of p , in other words, we assume that for all formulas φ , the implication $K_a\varphi \rightarrow \varphi$ holds.

It is often argued (indeed, already by Hintikka in [12]) that a distinguishing feature between knowledge and belief is that whereas knowledge is veridical, belief need not be, i.e., the scheme $B_ap \rightarrow p$ should not be assumed as an axiom for belief. This then simply entails that epistemic logics have veridicality as an axiom, and doxastic logics have not. Semantically speaking: the accessibility relations denoting knowledge are reflexive, those denoting belief need not be. But how then to deal with a situation where we want to express that “currently, a ’s beliefs happen to be true”? If we add $B_ap \rightarrow p$ as an axiom to our logic, the effect is that in all models (with respect to which the logic is complete), and in all states, all instances of that axiom are true, i.e., for all models M , for all states s and for all formulas φ , we then have $M, s \models B_a\varphi \rightarrow \varphi$. Given a model M and a state s we can express that a ’s belief that an individual proposition q holds is correct: $M, s \models B_aq \wedge q$. And we can express that a ’s belief about q is correct: $M, s \models (B_aq \rightarrow q) \wedge (B_a\neg q \rightarrow \neg q)$. But what we *cannot* express in modal logic is that $B_a\varphi \rightarrow \varphi$ holds for all φ in one state, without claiming at the same time it should hold throughout the model. As a consequence, we cannot express in the object language that agent b thinks that agent a ’s beliefs are correct, while agent c believes that a is wrong about a proposition q . The closest one gets to expressing that would be to say that for all φ , in M, s we have $M, s \models B_b(B_a\varphi \rightarrow \varphi) \wedge B_c((B_aq \wedge \neg q) \vee (B_a\neg q \wedge q))$ (but here, the quantification over φ is on a meta-level, and not in the scope of agent b). Neither can we say, in a temporal doxastic context, that a ’s beliefs now are correct, but tomorrow they need not be.

To give another example of the same phenomenon, suppose one adds the scheme $K_ap \rightarrow K_bp$ to a modal logic (b knows everything that a knows). Semantically, this means $R_b \subseteq R_a$. If the logic is about a set of agents A , then it becomes common knowledge among A that b knows at least what a knows! And if there is a notion of time, we have that it will always be the case that b knows at least what a knows, and, when having modalities for actions, it follows that no action can make it come about that a has a secret for b , in particular, it is impossible to inform a about something that b does not already know—this rules out dynamics that is, in contrast, very possible in DEL.

So, the general picture in modal logic that we take as our starting point is the following. One has a modal logic to which one adds an axiom scheme θ (say, $B_ap \rightarrow p$). If one is lucky, the scheme corresponds to a relational property $\Theta(x)$ (in the case above, Rxx). However, adding θ to the logic means having $\Theta(x)$ true everywhere, implying that θ is always true. What we are after is looking at ways to enforce the scheme θ *locally*. To do so, we will add a marker \Box to the modal language, such that \Box is true locally, in a state s , if and only if Θ is true, locally (i.e., Rss holds).

Doing so, we generalise work of [20], where a case study, in the context of a multi-agent logic **S5**, is given for ‘knowing at least as much as’, i.e., in our terminology, $\Box(a, b)$ in [20] equals $a \succeq b$, and our $\Theta(a, b)(x)$ is the property $\forall y(R_bxy \Rightarrow R_axy)$ in [20].

We will not only generalise the result of [20] to arbitrary

modal logics $\mathbf{K}(+\varphi_1, \dots, +\varphi_n)$ where φ_i are canonical axioms, but also we allow to add several markers at the same time. This then enables that we cannot only make global properties locally true, but it allows for far more subtle quantifications over formulas than is allowed in modal logic, enabling us to express properties like “If all of John’s beliefs are correct, then so must Mary’s beliefs be”, or “If John knows now everything that Mary knows, then that must have been true yesterday as well” or “If John’s beliefs are correct, then he must know that Mary’s beliefs are correct as well” (for more examples of such quantification, see Section 2.1).

This paper is organised as follows. In Section 1.1 we sketch how our machinery will look like. Then, in Section 2 we formally introduce three languages and present an example. Section 3 provides an axiomatisation of our extended modal logic, we come back to the example and make a case for completeness. Finally, in Section 4 we summarise and conclude.

1.1 To a Modal Logic with Local Schemes

In this section we introduce three languages to reason about Kripke models. The place where these languages meet are important for our set-up. Let us outline the overall approach at the hand of an example: formal definitions follow later in this section. First of all, we are interested in a *modal scheme* $\theta(a, b, p) = [a]p \rightarrow [b]p$ in a modal language \mathcal{L} (generally, we write $[a]\varphi$ for modal formulas, but for epistemic interpretations we may write $K_a\varphi$, and for doxastic ones $B_a\varphi$). To the modal language we add a relational atom $\Box(a, b)$, or, in this specific case $Sup(a, b)$, which will be true in a state s iff $\forall y(R_bsy \Rightarrow R_asy)$ holds. The latter property is a formula $\Theta(a, b)(s)$ in a first-order language \mathcal{L}^1 . Our modal logic should now formalise the idea that $\theta(a, b, c)$ and $\Box(a, b)$ ‘capture the same’. Rather than saying that the two are equivalent, the logic will take care that something along the following lines holds: consistency of a formula φ with an occurrence of $\neg\Box(a, b)$ is the same as consistency of φ with the occurrence of $\neg\Box(a, b)$ replaced by $\neg\theta(a, b, p)$ (if p is a fresh atom). For completeness of the logic, we then take care that in its canonical model, the truth of $\theta(a, b, p)$ in a specific world (i.e., maximal consistent set Δ) coincides with property $\Theta(\Delta)$. We show that our construction works because the second order formula $\forall P(\forall x(R_axy \Rightarrow Py) \Rightarrow \forall y(R_bxy \Rightarrow Py)) = \forall P\Theta(a, b, P)(x)$ is equivalent to $\Theta(x)$. The formula $\forall P\Theta(a, b, P)$ is an example of a formula from the third language that we use, i.e., a second-order language \mathcal{L}^2 .

The languages that we define are simple extensions of languages usually studied in standard modal logic [3, 2]. More specifically, our modal logic extends that of modal logic with some relational atoms \Box , the first order language is the standard language to reason about properties of accessibility relations, and the second order language is similar to the one usually obtained by applying the so-called standard translation to modal formulas. Our completeness proof, in turn, is an extension of ‘standard’ completeness proofs in modal logic: we sometimes have to add fresh atoms p to ensure that $\theta(a, b, p)$ is satisfied. However, we have borrowed ideas from [5] to prove our Extension Lemma 2 and ideas from [16] to make this lemma work ‘everywhere in the canonical model’. Space does not allow to include the proofs themselves, but we will make an effort to explain the overall idea and the construction of the canonical model.

2. LANGUAGE AND SEMANTICS

As outlined above, we deal with three languages, which are all interpreted over the same objects, i.e., Kripke models. The languages are an extended modal language \mathcal{L} , a first order language \mathcal{L}^1 and a second order language \mathcal{L}^2 .

For all languages, we assume a set of modality labels $A = \{a_1, \dots, a_{|A|}\}$. In the modal language, these will give rise to modalities $[a]$, and in the other two languages, we assume to have a binary relation R_a for each $a \in A$. For the latter two languages we also assume to have a set of variables $\mathcal{X} = \{x, y, \dots\}$. The variables will range over possible worlds: note that neither in \mathcal{L}^1 nor in \mathcal{L}^2 we assume to have constants. For \mathcal{L}^2 , we furthermore use a set $\Pi = \{P, P_1, P_2, \dots, Q, Q_1, Q_2, \dots\}$ of unary predicates. For each such predicate P in Π we assume to have an atomic proposition $p \in \pi$ that are building blocks for the modal language \mathcal{L} . On top of this, for this modal language \mathcal{L} we assume a finite set $\rho = \{\Box_1, \Box_2, \dots, \Box_m\}$ of relational atoms: they are nothing else than syntactic atoms of which the truth depends on local properties of accessibility relations (see the function I in Definition 1). Therefore, we will often write $\Box(a_1, \dots, a_n)$ rather than \Box to make this dependence clear, and treat \Box as if it were an n -ary relational predicate (rather than an atomic symbol). Our languages will be denoted $\mathcal{L}(A, \pi, \rho)$ (the modal language), $\mathcal{L}^1(A, \mathcal{X})$ (the first order language) and $\mathcal{L}^2(A, \Pi, \mathcal{X})$ (the second order language). If the parameters for the languages are clear, we will also write \mathcal{L} , \mathcal{L}^1 and \mathcal{L}^2 , respectively.

DEFINITION 1 (MODAL LANGUAGE). Let the sets A, π , and ρ be as described above. The modal language $\mathcal{L}(A, \pi, \rho)$ is defined as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid [a]\varphi \mid \Box(a_1, \dots, a_n)$$

where $a, a_1, \dots, a_n \in A$, $p \in \pi$ and \Box is an n -ary relational atom in ρ . Formula $\langle a \rangle\varphi$ is shorthand for $\neg[a]\neg\varphi$ and we also assume the usual definitions for disjunction, implication and bi-implication. If the modality is an epistemic one, the labels are agents, and we write $K_a\varphi$ rather than $[a]\varphi$. For a doxastic interpretation we write $B_a\varphi$, etc.

A formula without occurrences of relational atoms is called a purely modal formula. Suppose we have a multi-modal formula $\theta(a_1, \dots, a_n, p_1, \dots, p_k)$ where a_1, \dots, a_n are labels of modalities $[a_1], \dots, [a_n]$ and p_1, \dots, p_k are atoms. We will write \vec{a} for the tuple a_1, \dots, a_n and \vec{p} for p_1, \dots, p_k . When we write $a \in \vec{a}$ we mean that a is one of the labels occurring in the tuple \vec{a} , likewise for p and \vec{p} . Finally, for any tuple $\vec{x} = x_1, \dots, x_n$ with each x_i taken from some set X , we will write $\vec{x} \in \vec{X}$.

DEFINITION 2 (FIRST AND SECOND ORDER LANGUAGE). Let A and \mathcal{X} be given. First define a language $\mathcal{L}^+(A, \mathcal{X})$:

$$\Theta ::= R_a xy \mid \forall y \Theta \mid \neg\Theta \mid \Theta \ \& \ \Theta$$

with $a \in A$, and $x, y \in \mathcal{X}$. Now, our first order language $\mathcal{L}^1(A, \mathcal{X})$ is the one-free-variable sublanguage of \mathcal{L}^+ , i.e., the sublanguage of \mathcal{L}^+ consisting of all formulas with at most one variable not in the scope of a quantifier. If $\Theta \in \mathcal{L}^1(A, \mathcal{X})$ has x as its only free variable, and if a_1, \dots, a_n are all the modality labels occurring in Θ , we will also write $\Theta(\vec{a})(x)$ for Θ .

Finally, given A, Π and \mathcal{X} we define the second order language $\mathcal{L}^2(A, \Pi, \mathcal{X})$ as the one-free-variable fragment of

$$\hat{\Theta} ::= P(x) \mid R_a xy \mid \forall y \hat{\Theta} \mid \forall P \hat{\Theta} \mid \neg\hat{\Theta} \mid \hat{\Theta} \ \& \ \hat{\Theta}$$

with $P \in \Pi$, $x, y \in \mathcal{X}$ and $a \in A$. In $\mathcal{L}^1(A, \mathcal{X})$ and $\mathcal{L}^2(A, \Pi, \mathcal{X})$, existential quantification (using \exists) and implication (using \Rightarrow) are defined in a standard way.

We write Px for $P(x)$. As mentioned earlier, all languages will be interpreted over Kripke models.

DEFINITION 3 (KRIPKE MODELS AND FRAMES). Given A , π and ρ , a Kripke model is a tuple $M = \langle W, R, I, V \rangle$ where

- W is a set of possible worlds;
- $R : A \rightarrow \wp(W \times W)$ assigns a binary relation to each modality label
- $I : \rho \rightarrow \mathcal{L}^1(A, \mathcal{X})$ assigns a first order property to each relational atom in ρ
- $V : \pi \rightarrow \wp(W)$ assigns a set of possible worlds to each propositional variable

Rather than $(w, v) \in R(a)$ we will write $R_a wv$. A Kripke frame is a tuple $F = \langle W, R, I \rangle$ such that $\langle M, V \rangle = \langle W, R, I, V \rangle$ is a model. The ‘arity’ of a symbol $\Box \in \rho$ can be read off from its interpretation $I(\Box)$: if $I(\Box)$ refers to modalities a_1, \dots, a_n , then we may write $\Box(\vec{a})$ for \Box .

DEFINITION 4 (SEMANTICS OF MODAL FORMULAS). Let A and π be given. Also, let $M = \langle W, R, I, V \rangle$. Then we define, for $\varphi \in \mathcal{L}(A, \pi, \rho)$:

$$\begin{aligned} M, w \models p & \quad \text{iff} \quad w \in V(p) \\ M, w \models \neg\varphi & \quad \text{iff} \quad M, w \not\models \varphi \\ M, w \models \varphi \wedge \psi & \quad \text{iff} \quad M, w \models \varphi \text{ and } M, w \models \psi \\ M, w \models [a]\varphi & \quad \text{iff} \quad \text{for all } v \text{ if } R_a wv, \text{ then } M, v \models \varphi \\ M, w \models \Box(\vec{a}) & \quad \text{iff} \quad I(\Box(\vec{a}))(w) \text{ holds} \end{aligned}$$

The class of all models is denoted $\mathcal{K}(A, \pi, \rho)$. All models with interpretation I are denoted $\mathcal{K}(A, \pi, \rho, I)$. Validity in a model M is defined as usual. Moreover, $\mathcal{K}(A, \pi, \rho) \models \varphi$ means that for all I , $M = \langle W, R, I, V \rangle$, and all $w \in W$, we have $M, w \models \varphi$. Given I , we say that φ is I -satisfiable, if there is a model $M = \langle W, R, I, V \rangle$ and a $w \in W$ such that $M, w \models \varphi$. Formula φ is I -valid if $\neg\varphi$ is not I -satisfiable. If $F = \langle W, R, I \rangle$ is a frame, $F, w \models \varphi$ is defined as: for all valuations V , $\langle W, R, I, V \rangle, w \models \varphi$.

Interpretation of $\mathcal{L}^1(A, \mathcal{X})$ -formulas in a model $M = \langle W, R, I, V \rangle$ is straightforward. For $\mathcal{L}^2(A, \Pi, \mathcal{X})$, we assume that Ps holds for a predicate P iff $s \in V(p)$. In other words, the link between a propositional atom and a unary predicate is implicit by using lower-case and upper-case notation.

EXAMPLE 1. Let $\Box(a, b)$ be such that in M with interpretation I , we have $I(\Box(a, b)) = \Theta(a, b)$ where $\Theta(a, b)(x) = \forall y (R_b xy \Rightarrow R_a xy)$, saying that in the current world w , the set of a -successors of w is a superset of the set of b -successors of w . If this is the interpretation of $\Box(a, b)$, we will also write $\text{Sup}(a, b)$. As a second example, take $\Box = \Box(a)$ to be such that $I(\Box(a))(x) = R_a xx$. Note that $B_a \Box(a)$ can hence be interpreted as ‘ a believes that his beliefs are correct’, since $M, w \models B_a \Box(a)$ does entail that for all φ , $M, w \models$

Table 1: In this table, \vec{a} is a sequence a or (a, b) or (a, b, c) of modality labels, and \vec{p} is either the single atom p or the sequence p, q . $\Theta(\vec{a})(x)$ is a property of a state x , and $\Box(\vec{a})$ is a name in the object language such that $\Box(\vec{a})$ holds at w iff $\Theta(\vec{a})(w)$ holds of M .

$\theta(\vec{a}, \vec{p})$	$\Theta(\vec{a})(x)$	$\Box(\vec{a})$
$[a]p \rightarrow [b]p$	$\forall y(R_bxy \Rightarrow R_axy)$	$Sup(a, b)$
$[c]p \rightarrow [a][b]p$	$\forall y, z((R_axy \& R_byz) \Rightarrow R_cxz)$	$Trans(a, b, c)$
$\neg[a]\perp$	$\exists y R_axy$	$Ser(a)$
$[a]p \rightarrow p$	R_axx	$Refl(a)$
$\neg[a]p \rightarrow [b]\langle c \rangle p$	$\forall yz((R_axy \& R_bxz) \Rightarrow R_cyz)$	$Eucl(a, b, c)$
$\langle a \rangle p \rightarrow \langle b \rangle \langle c \rangle p$	$\forall z(R_axz \Rightarrow \exists y R_bxy \& R_cyz)$	$Dens(a, b, c)$

Table 2: For every modal formula $\theta(\vec{a}, \vec{p})$ from Table 1, we give the second order translation $\hat{\Theta}(\vec{a}, \vec{P})(x)$.

$\theta(\vec{a}, \vec{p})$	$\hat{\Theta}(\vec{a}, \vec{P})(x)$
$[a]p \rightarrow [b]p$	$\forall y(R_bxy \Rightarrow Py) \Rightarrow \forall z(R_axz \Rightarrow Pz)$
$[c]p \rightarrow [a][b]p$	$\forall w(R_cxw \Rightarrow Pw) \Rightarrow \forall y(R_axy \Rightarrow \forall z(R_byz \Rightarrow Pz))$
$\neg[a]\perp$	$\neg \forall y(R_axy \Rightarrow \perp)$
$[a]p \rightarrow p$	$\forall y(R_axy \Rightarrow Py) \Rightarrow Px$
$\neg[a]p \rightarrow [b]\langle c \rangle p$	$\neg \forall w(R_axw \Rightarrow Pw) \Rightarrow \forall y(R_bxy \Rightarrow \exists z(R_cyz \& Pz))$
$\langle a \rangle p \rightarrow \langle b \rangle \langle c \rangle p$	$\exists w(R_axw \& Pw) \Rightarrow \exists y(R_bxy \& \exists z(R_cyz \& Pz))$

$B_a(B_a\varphi \rightarrow \varphi)$ (but see Remark 1). As a final example, take $\Box(a, b, c)$ with $I(\Box(a, b, c))(x) = \forall y \forall z((R_axy \& R_byz) \Rightarrow R_cxz)$ we will write $Trans(a, b, c)$ for $\Theta(a, b, c)$. Of course, a special case of this is $\Box = \Box(a, a, a)$ saying that currently, at world w , the relation R_a is transitive. For more examples, see Table 1.

REMARK 1. Take $\Box(a)$ and M such that $I(\Box(a)) = \forall x R_axx$. Note that although $M, w \models \Box(a)$ entails that agent a 's beliefs are correct, the converse is not true, as the following example shows. Let $M = \langle W, R, I, V \rangle$ be such that $W = \{w, v\}$, and $R_a = \{(w, u), (u, w)\}$. Moreover, assume that for all p , $w \in V(p)$ iff $u \in V(p)$. Since M, w and M, v are bisimilar [2, Chapters 1 and 5] models, we have $M, w \models \varphi$ iff $M, u \models \varphi$, and hence $M, w \models B_a\varphi \rightarrow \varphi$, for all purely modal φ . However, since $(w, w) \notin R_a$, we have $M, w \models \neg \Box(a)$.

Note that, since $\Theta(\vec{a})(w)$ does not refer to atomic propositions p (or, rather predicates P), we have that $\Theta(\vec{a})(w)$ holds in the model M iff $\Theta(\vec{a})(w)$ holds in the frame F .

DEFINITION 5 (STANDARD TRANSLATION). Fix sets A , π , Π , and ρ . Fix an interpretation I and write $I(\Box(\vec{a})) = \Theta_{\Box}(\vec{a})$. We define $ST_I : \mathcal{L}(A, \pi, \rho) \times \mathcal{X} \rightarrow \mathcal{L}^2(A, \Pi, \mathcal{X})$ by

$$\begin{aligned}
ST_I(p)(x) &= P(x) \\
ST_I(\Box \vec{a})(x) &= \Theta_{\Box}(\vec{a})(x) \\
ST_I(\neg \varphi)(x) &= \neg ST_I(\varphi)(x) \\
ST_I(\varphi \wedge \psi)(x) &= ST_I(\varphi)(x) \& ST_I(\psi)(x) \\
ST_I([a]\varphi)(x) &= \forall y(R_axy \Rightarrow ST_I(\varphi)(y))
\end{aligned}$$

In the last clause, y is assumed to be a fresh variable. If φ is purely modal (i.e., φ is \Box -free), $ST_I(\varphi)$ does not depend on the interpretation I and we write $ST(\varphi)$ in such a case.

Note that the standard translation $ST(\theta(\vec{a}, \vec{p}))$ of a modal formula involving modalities \vec{a} and atoms \vec{p} is typically a formula $\hat{\Theta}(\vec{a}, \vec{P})(x)$ involving binary relations R_a (one for

each $a \in \vec{a}$) and predicates P (one for each $p \in \vec{p}$) and a free variable x .

EXAMPLE 2. Take $\theta(a, b, p) = [a]p \rightarrow [b]p$. Then we have that $ST_I(\theta(a, b, p))(x) =$

$$\forall y(R_bxy \Rightarrow Py) \Rightarrow \forall z(R_axz \Rightarrow Pz)$$

If $\theta(a, b, c, p) = [c]p \rightarrow [a][b]p$, we have $ST_I(\theta(a, b, c, p)) =$

$$\forall w(R_cxw \Rightarrow Pw) \Rightarrow \forall y(R_axy \Rightarrow \forall z(R_byz \Rightarrow Pz))$$

The following is straightforward from classical modal theory, except for the case of $\theta(\vec{a}, \vec{p}) = \Box(\vec{a})$, in which case it follows directly from the truth definition (Definition 4) for \Box -formulas.

LEMMA 1. Let I be an interpretation of relational symbols and let $\theta(\vec{a}, \vec{p})$ be a modal formula. Let $\hat{\Theta}(\vec{a}, \vec{P})(x)$ be its second order translation $ST_I(\theta(\vec{a}, \vec{p}))$. Then, for all models $M = \langle W, R, I, V \rangle$, all frames $F = \langle W, R, I \rangle$ and worlds $w \in W$ we have

1. $M, w \models \theta(\vec{a}, \vec{p})$ iff $\hat{\Theta}(\vec{a}, \vec{P})(w)$ holds in M
2. $M \models \theta(\vec{a}, \vec{p})$ iff $\forall x \hat{\Theta}(\vec{a}, \vec{P})(x)$ holds in M
3. $F, w \models \theta(\vec{a}, \vec{p})$ iff $\forall \vec{P} \hat{\Theta}(\vec{a}, \vec{P})(w)$ holds in F
4. $F \models \theta(\vec{a}, \vec{p})$ iff $\forall x \forall \vec{P} \hat{\Theta}(\vec{a}, \vec{P})(x)$ holds in F

Table 2 provides the standard translation $\hat{\Theta}(\vec{a}, \vec{P})(x)$ of the modal formulas $\theta(\vec{a}, \vec{p})$ that we introduced in Table 1.

DEFINITION 6. Let $\theta(\vec{a}, \vec{p})$ be a purely modal formula from $\mathcal{L}(A, \pi, \rho)$ and suppose that $\Theta \in \mathcal{L}^1(A, \mathcal{X})$ is such that $\Theta(\vec{a})(x)$ is equivalent with the second order formula $\forall \vec{P} \hat{\Theta}(\vec{a}, \vec{P})(x)$ where $\hat{\Theta}(\vec{a}, \vec{P})(x) = ST(\theta(\vec{a}, \vec{p}))(x)$. Then we say that $\theta(\vec{a}, \vec{p})$ characterises $\Theta(\vec{a})(x)$.

If $\theta(\vec{a}, \vec{p})$ characterises $\Theta(\vec{a})(x)$ then we have that $F, w \models \theta(\vec{a}, \vec{p})$ if $\Theta(\vec{a})(w)$ holds. In other words, $\theta(\vec{a}, \vec{p})$ corresponds with $\Theta(\vec{a})$. There are many well known classes of modal formulas $\theta(\vec{a}, \vec{p})(x)$ for which it is guaranteed that the second order formula $\forall \vec{PST}(\theta(\vec{a}, \vec{p}))$ is equivalent to a formula $\Theta(\vec{a})(x) \in \mathcal{L}^1(A, \mathcal{K})$. A large set of formulas for which this is true is the set of so-called *Sahlqvist formulas*. Moreover, given such a Sahlqvist formula $\theta(\vec{a}, \vec{p})$, its first order equivalent $\Theta(\vec{a})(x)$ can be effectively computed from it [3, Theorem 3.54]. So for Sahlqvist formulas, we can effectively find the first-order formula that it characterises. All the formulas $\theta(\vec{a}, \vec{p})$ from Table 1 are (equivalent to) Sahlqvist formulas.

Take the specific example in a doxastic context where $\Theta(a)(x)$ is $\forall x R_a x x$, and $I(\Box(a)) = \Theta(a)$, note that $\theta(a, p) = (B_a p \rightarrow p)$ characterises $\Theta(a)(x)$ but still, as shown in Remark 1, the formulas $\Box(a)$ and $\theta(a, p)$ are not *equivalent*. Still, the two should be strongly connected, in a sense we will explain in Section 3. We first look at an example, involving our extended modal language.

2.1 A Simple Example

Consider five friends, Joey, Chandler, Ross, Monica and Phoebe (or j, c, r, m and p , for short). In this example, we use ‘think’ and ‘believe’ for the same thing. Joey believes that Monica’s beliefs are at least as accurate as Ross’ beliefs, i.e., Joey believes that if Ross’ beliefs are correct, so must Monica’s be (A). Joey also believes that Monica thinks that Chandler believes anything that Monica believes (B). Although Joey does not think that he believes everything he knows (he thinks that he knows he cannot find a job as an actor, but at the same time cannot believe it), he actually believes anything he knows (C). Moreover, Joey thinks that Chandler’s beliefs are consistent (D). Finally, Joey happens to know that Monica believes that Phoebe is in competition with her for Chandler’s attention, but at the same time Joey thinks that Chandler believes that Phoebe is not in competition with Monica for his attention (E). Then, we conclude that Joey believes that Ross’ beliefs are not guaranteed to be correct (F), or, better, that Joey believes he may assume that some formula is believed by Ross, but not true (F').

We first give a (semi-formal) formalisation of our assumption using a modal logic that allows for quantification over formulas. Let z represent the proposition that Joey cannot find himself a job as an actor, and let q be the proposition that Phoebe is in competition with Monica for Chandler’s attention. This formalisation is given in Table 3, where assumption A in the episode is represented as a , etc. The formalisation in our language $\mathcal{L}(A, \pi, \rho)$ follows in Table 4.

We can now be more precise about what it means that our language can do more than just formalising a local version of a global property. For instance, the global property $B_a p \rightarrow p$ will have a local counterpart $Refl(a)$. Locally, this will denote something that is similar to $\forall \varphi (B \varphi \rightarrow \varphi)$. But if one looks at the ‘translation’ a in Table 3 of the assumption A above, i.e., $B_j(\forall \varphi (B_r \varphi \rightarrow \varphi) \rightarrow \forall \varphi (B_m \varphi \rightarrow \varphi))$, it becomes clear that this is different from the quantification $g : \forall \varphi B_j((B_r \varphi \rightarrow \varphi) \rightarrow (B_m \varphi \rightarrow \varphi))$, which one would get as a local counterpart of an axiom $B_j((B_r p \rightarrow p) \rightarrow (B_m p \rightarrow p))$. That a and g are not equivalent, can be seen in the model M, w of Figure 1, where a is true in M, w , but g is not: for the latter, $\varphi = p$ provides a counterexample. That a is true in M, w is easily seen from realising that a is formalised by a' .

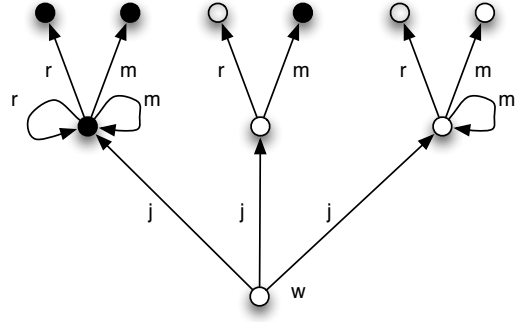


Figure 1: A model M, w . The atom p is true exactly in the worlds that are filled black.

Table 3: A semi-formal translation of the episode

a	$B_j(\forall \varphi (B_r \varphi \rightarrow \varphi) \rightarrow \forall \varphi (B_m \varphi \rightarrow \varphi))$
b	$B_j B_m(\forall \varphi (B_m \varphi \rightarrow B_c \varphi))$
c	$\neg B_j(K_j z \rightarrow B_j z) \wedge \forall \varphi (K_j \varphi \rightarrow B_j \varphi)$
d	$B_j \forall \varphi (\neg (B_c \varphi \wedge B_c \neg \varphi))$
e	$K_j B_m q \wedge B_j B_c \neg q$
f	$B_j \neg \forall \varphi (B_r \varphi \rightarrow \varphi)$

We then formalise the same episode using the relational atoms $\Box(\vec{a})$ introduced in Table 1, which results in Table 4. Abusing the language somewhat, we write $Sup(kj, j)$ for the relational atom corresponding to $K_j \varphi \rightarrow B_j \varphi$ —from a language point of view, K_j and B_j are simply two different modal operators, say $[i]$ and $[k]$.

Table 4: A formalisation of the episode

a'	$B_j(Refl(r) \rightarrow Refl(m))$
b'	$B_j B_m Sup(m, c)$
c'	$\neg B_j(K_j z \rightarrow B_j z) \wedge Sup(kj, k)$
d'	$B_j Ser(c)$
e'	$K_j B_m q \wedge B_j B_c \neg q$
f'	$B_j \neg Refl(r)$

3. AXIOMATIZATION

The aim of this section is to provide an axiomatisation for modal logics that are enriched with some relational atoms $\Box_1(\vec{a}_1), \dots, \Box_m(\vec{a}_m)$, such that for every $\Box_k(k \leq m)$, there is a modal formula $\theta_{\Box_k}(\vec{a}, \vec{p})$ such that, at least on frames, the two ‘mean the same thing’. In fact, the logic $\mathbf{K}(A, \pi, \rho, I)$ that we define should be sound and complete with respect to $\mathcal{K}(A, \pi, \rho, I)$, so our aim for our logic is that for all formulas $\varphi \in \mathcal{L}(A, \pi, \rho)$, the notions $\mathbf{K}(A, \pi, \rho, I) \vdash \varphi$ and $\mathcal{K}(A, \pi, \rho, I) \models \varphi$ coincide. The idea to achieve this is as follows. First of all, suppose that for every relational atom $\Box(\vec{a})$ and fixed interpretation I we have a formula $\theta_{\Box}(\vec{a}, \vec{p})$ such that $\theta_{\Box}(\vec{a}, \vec{p})$ characterises $I(\Box(\vec{a}))$. Then, for each $\Box(\vec{a})$ and related $\theta_{\Box}(\vec{a}, \vec{p})$ we add an axiom $\Box(\vec{a}) \rightarrow \theta_{\Box}(\vec{a}, \vec{p})$ to our logic $\mathbf{K}(A, \pi, \rho, I)$.

Adding the other direction as an implication does not work, as the example $\Box(\vec{a}) = \Box(a, b) = \text{Sup}(a, b)$ and $\theta(\vec{a}, \vec{p}) = [a]p \rightarrow [b]p$ shows: $([a]p \rightarrow [b]p) \rightarrow \text{Sup}(a, b)$ is not a validity: the antecedent may be true due to some specific choice of p . However, what we semantically *do* have is the following: suppose that we have $\mathcal{K}(A, \pi, \rho, I) \models \varphi \rightarrow ([a]p \rightarrow [b]p)$, where p does not occur in φ . This then means that φ must entail that (locally) all b successors are a successors, i.e., $\mathcal{K}(A, \pi, \rho, I) \models \varphi \rightarrow \text{Sup}(a, b)$, because if the latter would not hold, there would be a model $M = \langle W, R, I, V \rangle$ such that $M, w \models \varphi \wedge \neg \text{Sup}(a, b)$. But since p does not occur in φ , we can change the valuation V for p freely without changing that of φ , in particular we can choose V' such that $x \in V'(p) \leftrightarrow R_ax$ (and $V'(q) = V(q)$ for atoms $q \neq p$). It is easy to see that in the resulting model $M' = \langle W, R, I, V' \rangle$ we have $M', w \models \varphi \wedge \neg([a]p \rightarrow [b]p)$: a contradiction. This means that we need to be able to infer the following in $\mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I)$:

$$\begin{aligned} \text{If } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow ([a]p \rightarrow [b]p) \quad (1) \\ \text{then } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow \text{Sup}(a, b), \\ \text{where } p \notin \varphi \end{aligned}$$

The rule (1) can be understood as follows. If p does not occur in φ , and $\varphi \rightarrow ([a]p \rightarrow [b]p)$ is true at a state s , then φ must carry sufficient information such that $[a]p \rightarrow [b]p$ must hold (it will not be because of specific requirements on p imposed by φ) and hence we must have $\varphi \rightarrow \text{Sup}(a, b)$ holding at s as well. But in fact we can do the same reasoning that involves *successors* of s : Suppose φ implies that in all R_c successors t of s , we have $M, t \models [a]p \rightarrow [b]p$. Then (in the same way as for s), we must have $M, t \models \varphi \rightarrow \text{Sup}(a, b)$. In other words, the following should hold for $\mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I)$:

$$\begin{aligned} \text{If } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow [c]([a]p \rightarrow [b]p) \quad (2) \\ \text{then } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow [c]\text{Sup}(a, b), \\ \text{where } p \notin \varphi \end{aligned}$$

And the same should hold for all R_d successors u of all R_c -successors t of s , i.e., we also have a valid rule if we replace $[c]$ in (2) by $[c][d]$. But also, we have the following. Suppose that p does not occur in φ or ψ . If $M, s \models \varphi \rightarrow [c](\psi \rightarrow ([a]p \rightarrow [b]p))$, it means that if φ is true in s , then in all R_c successors t of s we have $M, t \models \psi \rightarrow ([a]p \rightarrow [b]p)$, and we have argued above that we then should also have $M, t \models \psi \rightarrow \text{Sup}(a, b)$. I.e., we have:

$$\begin{aligned} \text{If } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow [c](\psi \rightarrow ([a]p \rightarrow [b]p)) \quad (3) \\ \text{then } \mathcal{K}(A, \pi, \{\text{Sup}(a, b)\}, I) \models \varphi \rightarrow [c](\psi \rightarrow \text{Sup}(a, b)), \\ \text{where } p \notin \varphi, \psi \end{aligned}$$

To formalise that a property $\theta_{\Box}(\vec{a}, \vec{p})$ holds after arbitrary sequences $\varphi_1 \rightarrow [a_1](\varphi_2 \rightarrow \dots [a_{n-1}](\varphi_n \rightarrow \theta_{\Box}(\vec{a}, \vec{p})) \dots)$, we follow [20] and introduce *pseudo modalities*: we will then present an inference rule \mathbf{R}_{\Box} for every $\Box \in \rho$ to our axiomatisation $\mathbf{K}(A, \pi, \rho, I)$.

DEFINITION 7 (PSEUDO MODALITIES). *We define the following pseudo modalities, which are (possibly empty) sequences $s = ()$ or $s = (s_1, \dots, s_n)$, where each s_i is a formula or a modality label. The formula $\langle s \rangle \varphi$ represents an $\mathcal{L}(A, \Pi, \rho)$ formula, as follows:*

$$\begin{aligned} \langle () \rangle \varphi &= \varphi \\ \langle \psi, s_2, \dots, s_n \rangle \varphi &= \psi \wedge \langle s_2, \dots, s_n \rangle \varphi \\ \langle a, s_2, \dots, s_n \rangle \varphi &= \langle a \rangle (\langle s_2, \dots, s_n \rangle \varphi) \end{aligned}$$

We also define $[s]\varphi$ as $\neg \langle s \rangle \neg \varphi$. We say that \vec{p} does not occur in s (and write $\vec{p} \notin s$) if none of the atoms p occurring in \vec{p} does occur in any of the formulas s_i in s .

So, for instance $\langle a, \psi, b \rangle \varphi$ is an abbreviation of $\langle a \rangle (\psi \wedge \langle b \rangle \varphi)$, while $[a, \psi, b]\varphi$ is $[a](\psi \rightarrow [b]\varphi)$.

DEFINITION 8 (PROOF SYSTEM). *Fix A, π and ρ . Moreover, fix an interpretation $I : \rho \rightarrow \mathcal{L}^1(A, \mathcal{X})$ such that for every $\Box \in \rho$, there is a $\theta_{\Box}(\vec{a}, \vec{p})$ such that the modal formula $\theta_{\Box}(\vec{a}, \vec{p})$ characterises the first order formula $I(\Box)$. Then, the following comprises the axioms and inference rules of the logic $\mathbf{K}(A, \pi, \rho, I)$*

Prop All instances of propositional tautologies

$$\mathbf{K} \quad [a](\varphi \rightarrow \psi) \rightarrow ([a]\varphi \rightarrow [a]\psi)$$

$$\mathbf{Ax}_{\Box} \quad \Box(\vec{a}) \rightarrow \theta_{\Box}(\vec{a}, \vec{p})$$

MP From $\varphi \rightarrow \psi$ and φ , infer ψ

Nec From φ , infer $[a]\varphi$

R_□ From $\langle s \rangle \neg \theta_{\Box}(\vec{a}, \vec{p}) \rightarrow \varphi$, infer $\langle s \rangle \neg \Box(\vec{a}) \rightarrow \varphi$, where \vec{p} does not occur in φ or s .

US From φ infer $\varphi[\psi/p]$.

MP stands for *Modus Ponens*, **Nec** for *Necessitation*, and **US** for *Uniform Substitution* ($\varphi[\psi/p]$ stands for substitution of ψ for every occurrence of p in φ). If $\Box(\vec{a})$ and $\theta_{\Box}(\vec{a}, \vec{p})$ are connected through the axiom \mathbf{Ax}_{\Box} and inference rule \mathbf{R}_{\Box} , we say they are *axiomatically linked* (through axiom \mathbf{Ax}_{\Box} and rule \mathbf{R}_{\Box}). If there is a derivation of a formula φ from a set of formulas Γ using Γ and the axioms and inference rules from $\mathbf{K}(A, \pi, \rho, I)$ we write $\Gamma \vdash_{\mathbf{K}(A, \pi, \rho, I)} \varphi$, or $\Gamma \vdash_{\mathbf{K}} \varphi$, for short.

THEOREM 1 (SOUNDNESS). *For all $\varphi \in \mathcal{L}(A, \pi, \rho)$, if $\mathbf{K}(A, \pi, \rho, I) \vdash \varphi$ then $\mathcal{K}(A, \pi, \rho, I) \models \varphi$.*

3.1 Back to Our Example

To formalise the derivation of Table 4, let the set of modalities $A = \{c, j, m, p, r\}$, let $\pi = \{q, z\}$ and let $\rho = \{\text{Refl}(r), \text{Refl}(m), \text{Sup}(c, m), \text{Ser}(c), \text{Sup}(kj, j)\}$ and those atoms are axiomatically linked with their ‘natural’ modal counterparts (see Table 1 and for $\text{KimplB}(j)$ we take $K_j p \rightarrow B_j p$). Let the resulting logic be $\mathbf{K}(A, \pi, \rho, I)$.

First of all, from c' and $\mathbf{Ax}_{\text{KimplB}(j)}$ we derive $K_j B_m q \rightarrow B_j B_m q$. Together with e' this gives e'' : $B_j B_m q \wedge B_j B_c \neg q$. From d' , i.e., $B_j \text{Ser}(c)$ and $\mathbf{Ax}_{\text{Ser}(c)}$, we get $B_j(B_c \neg q \rightarrow \neg B_c q)$. Combining this with e'' gives $B_j B_m q \wedge B_j \neg B_c q$, which is equivalent to $B_j \neg(B_m q \rightarrow B_c q)$ (*).

From b' and $\mathbf{Ax}_{\text{Sup}(m, c)}$ we derive $B_j B_m(B_m p \rightarrow B_c p)$, for any p (**). Now, take the formula $\psi = (B_m q \rightarrow B_c q)$. From (*) we have $B_j \neg \psi$, and from (**) we conclude $B_j B_m \psi$. In other words, we found a formula ψ for which $B_j \neg(B_m \psi \rightarrow \psi)$. Now using the contrapositive of axiom $\mathbf{Ax}_{\text{Refl}(m)}$, we obtain $B_j \neg \text{Refl}(m)$, which is our conclusion f' .

Now one may wonder whether this also warrants the conclusion f , but as should be clear from Remark 1, $B_j \neg \text{Refl}(m)$ and $B_j \neg \forall \varphi (B_m \varphi \rightarrow \varphi)$ are not the same thing. However, what we *do* have is the following. Let φ be $a' \wedge b' \wedge c' \wedge d' \wedge e'$, and let s be B_j , then what we have proven now is

$$\mathbf{K}(A, \pi, \rho, I) \vdash \varphi \rightarrow B_j \neg \text{Refl}(m) \quad (4)$$

But from this, if not *derive*, we can safely *assume* that, given φ , there is some formula ψ for which $B_j(B_m\psi \wedge \neg\psi)$, because if this were not the case, we would have, for some atom p not occurring in φ :

$$\mathbf{K}(A, \pi, \rho, I) \vdash \varphi \rightarrow B_j(B_m p \rightarrow p) \quad (5)$$

From which, using rule $\mathbf{R}_{\square}(\vec{a})$, we would conclude

$$\mathbf{K}(A, \pi, \rho, I) \vdash \varphi \rightarrow B_j \text{Refl}(m) \quad (6)$$

which either means we have a derivation for $B_j \perp$ (Joey believes anything), or, if we assume the conjunct $g' = \text{Ser}(j)$ to be also part of φ (expressing, that actually, Joey's beliefs are consistent), that we have derived a contradiction with (4).

It is worth noting how the axiomatisation makes it possible that some relational atoms (and hence some first-order frame properties) only hold in the scope of a modal operator (like in property a' and b' for example): the axiom \mathbf{Ax}_{\square} and rule \mathbf{R}_{\square} do not require that some relational properties hold, they only specify what should be the case if they hold.

3.2 Completeness

DEFINITION 9. A theory Γ is a set of formulas. For π a set of propositional atoms, Γ is a π -theory if all propositional atoms in Γ are from π . Given a logic \mathbf{L} , a theory Γ is \mathbf{L} -consistent if \perp cannot be derived from Γ using the axioms and inference rules of \mathbf{L} . A theory Γ is a maximal \mathbf{L} -consistent π -theory if it is consistent and no π -theory Δ is \mathbf{L} -consistent while at the same time $\Gamma \subset \Delta$. For a logic $\mathbf{K}(A, \pi, \rho, I)$, a set of formulas Γ is a witnessed π -theory if for every $\langle s \rangle \neg \square(\vec{a}) \in \Gamma$, there are atoms \vec{p} such that $\langle s \rangle \neg \theta_{\square}(\vec{a}, \vec{p}) \in \Gamma$, where $\square(\vec{a})$ and $\theta_{\square}(\vec{a}, \vec{p})$ are axiomatically linked. If Γ is not witnessed, then a formula $\langle s \rangle \neg \square(\vec{a})$ for which there is no $\langle s \rangle \neg \theta_{\square}(\vec{a}, \vec{p}) \in \Gamma$, is called a defect for the theory Γ . Finally, Γ is said to be fully witnessed, if it is witnessed and for every formula of the form $\langle s \rangle \varphi$, either that formula or its negation is in Γ .

LEMMA 2 (EXTENSION LEMMA). Let Σ be a $\mathbf{K}(A, \pi, \rho, I)$ -consistent π -theory. Let $\pi' \supseteq \pi$ be an extension of π by a countable set of propositional variables. Then there is a maximal $\mathbf{K}(A, \pi', \rho, I)$ -consistent, witnessed π' -theory Σ' extending Σ .

Before we outline a proof, we first define some languages.

DEFINITION 10. Let the set of agents A , the set of atoms π and the set of relational atoms ρ be fixed. Let $\mathcal{L}(A, \pi, \rho)$ be as in Definition 1. Let $\pi^0 = \{p_0, p_1, \dots\}$ be a set of fresh atomic variables, i.e., $\pi \cap \pi^0 = \emptyset$ and let $\pi' = \pi \cup \pi^0$. Let $\pi_n = \pi \cup \{p_i \mid i \leq n\}$. Define \mathcal{L}_n to be $\mathcal{L}(A, \pi_n, \rho)$, and let \mathcal{L}_{ω} be $\mathcal{L}(A, \pi', \rho)$. A theory $\Delta \subseteq \Sigma$ is called an approximation if for some n it is a consistent π_n -theory. For such a theory, and any number k , the sequence $\vec{p} = \langle p_{n+1}, \dots, p_{n+k} \rangle$ is a new sequence \vec{p} for Δ if n is the least number such that Δ is a π_n -theory.

PROOF OF LEMMA 2 (SKETCH). Assume an enumeration of ψ_0, ψ_1, \dots of all formulas of the form $\langle s \rangle \neg \square(\vec{a})$, where s is a pseudo modality and $\square(\vec{a}) \in \rho$. Define

$$\Delta^+ = \begin{cases} \Delta \cup \{\langle s \rangle \neg \theta_{\square}(\vec{a}, \vec{p})\} & \text{where } \vec{p} \text{ is a new sequence} \\ & \text{for } \Delta, \text{ and } \langle s \rangle \neg \square(\vec{a}) \text{ is the} \\ & \text{first defect for } \Delta, \\ \Delta & \text{if this exists} \\ & \text{otherwise} \end{cases}$$

Clearly, by \mathbf{Ax}_{\square} , the set Δ^+ is consistent when Δ is and hence, if Δ is an approximation, so is Δ^+ . To define the extension Σ' of Σ , assume $\varphi_0, \varphi_1, \dots$ to be an enumeration of the formulas in \mathcal{L}_{ω} , and define $\Sigma_0 = \Sigma$, and

$$\begin{aligned} \Sigma_{2n+1} &= \begin{cases} \Sigma_{2n} \cup \{\varphi_n\} & \text{if this is consistent} \\ \Sigma_{2n} \cup \{\neg\varphi_n\} & \text{else} \end{cases} \\ \Sigma_{2n+2} &= (\Sigma_{2n+1})^+ \end{aligned}$$

Finally, let $\Sigma' = \bigcup_{n \in \omega} \Sigma_n$. By construction, $\Sigma' \supset \Sigma$ is a maximal $\mathbf{K}(A, \pi, \rho, I)$ -consistent, witnessed π' -theory. \square

DEFINITION 11 (CANONICAL MODEL). The canonical model $M^c = (W^c, R^c, I, V^c)$ for the logic $\mathbf{K}(A, \pi, \rho, I)$ has:

- $W^c = \{\Gamma \mid \Gamma \text{ is a maximal } \mathcal{L}_{\omega}\text{-consistent witnessed } \pi'\text{-theory}\};$
- $R_a^c \Gamma \Delta$ iff for all $\varphi \in \mathcal{L}_{\omega}$ it holds that if $[a]\varphi \in \Gamma$, then $\varphi \in \Delta$;
- I as given as a parameter of the logic;
- $V_p^c = \{\Gamma \mid p \in \pi' \cap \Gamma\}.$

LEMMA 3. Suppose the following holds:

1. $\theta_{\square}(\vec{a}, \vec{p})$ is a purely modal formula;
2. $\theta_{\square}(\vec{a}, \vec{p})$ and $\square(\vec{a})$ are linked through the rule R_{\square} and the axiom \mathbf{Ax}_{\square} ;
3. The first order formula $I(\square(\vec{a})) = \Theta(\vec{a})(x)$ is equivalent with the second order formula $\forall \vec{P} \hat{\Theta}(\vec{a}, \vec{P})(x)$ where $\hat{\Theta}(\vec{a}, \vec{P})(x)$ is $\text{ST}(\theta_{\square}(\vec{a}, \vec{p}))(x)$.

Then, in the canonical model, $\square(\vec{a})$ and $\Theta(\vec{a})(x)$ are connected as in Definition 4, i.e., for all $\Delta \in M^c$ we have $M^c, \Delta \models \square(\vec{a})$ iff in M^c it holds that $I(\square(\vec{a}))(\Delta)$.

Lemma 3 paves the way for a coincidence lemma that guarantees that membership and truth in the canonical model coincide. We then get:

THEOREM 2. If the assumptions of Lemma 3 hold, the logic $\mathbf{K}(A, \pi, \rho, I)$ is sound and complete with respect to the class of $\mathcal{K}(A, \pi, \rho, I)$ models.

DEFINITION 12. A purely modal formula φ is canonical for a first order property Φ , if the canonical model for the modal logic $(\mathbf{K} + \varphi)(A, \pi, \rho, I)$ has the property Φ .

There are many examples of canonical formulas: all Sahlqvist formulas are canonical [8].

THEOREM 3. Let φ_i be canonical for Φ_i , $i \leq n$. Then the logic $(\mathbf{K} + \varphi_1, \dots, \varphi_n)(A, \pi, \rho, I)$ is sound and complete with respect to all models in $\mathcal{K}(A, \pi, \rho, I)$ that satisfy Φ_1, \dots, Φ_n .

4. CONCLUSION

First, note that our modelling of locality is different from local frame correspondence as defined in [3], and quite distant from the use of *local propositions* in epistemic logic [6], propositions that do not change truth value within an agent's equivalence class.

We have so far assumed that the properties of $\square(\vec{a})$ are those specified by the axiom \mathbf{Ax}_{\square} and rule \mathbf{R}_{\square} . However, one

can add connections between $\Box(\vec{a})$ and modal formulas, or between different $\Box_1(\vec{a}_1)$ and $\Box_2(\vec{a}_2)$ atoms. For instance

$$\text{Refl}(a, a) \rightarrow \text{Trans}(a, a, a) \quad (7)$$

added to an epistemic logic has the effect that whenever a 's knowledge is veridical, a is also positively introspective. I.e., we would have, semantically, that whenever $M, s \models K_a \varphi \rightarrow \varphi$, for all φ , then also $M, s \models K_a \varphi \rightarrow K_a K_a \varphi$, for all φ . This again is a property that cannot be expressed in standard, 'global' modal logic. As a second example, in an epistemic temporal modal logic, one could add an axiom

$$\text{Trans}(a, a, a) \rightarrow F(\text{Trans}(a, a, a) \wedge \text{Eucl}(a, a, a)) \quad (8)$$

saying that whenever agent a is positively introspective, he will eventually also become negatively introspective. As a third example, a simple axiom like

$$\text{Ser}(a) \rightarrow \text{Ser}(b) \quad (9)$$

in a doxastic setting would mean that whenever a 's beliefs are consistent, those of b must be consistent as well.

It is possible to view some standard results in modal logic concerning completeness of modal systems as obtained as special cases from our local logic. If the conditions of Theorem 2 are satisfied, and one adds a $\Box(\vec{a})$ as an *axiom*, one immediately gets completeness with respect to the class of models that satisfy $I(\Box(\vec{a}))$. For instance, in a logic with axioms and rules for $\text{Refl}(a)$, adding $\text{Refl}(a)$ itself as an axiom gives a modal system that is sound and complete with respect to the class of reflexive Kripke models! Of course, this amounts to the same thing as adding $\theta_{\Box}(\vec{a}, \vec{p})$, as is directly clear from rule \mathbf{R}_{\Box} (take $\varphi = \perp$ and s the empty sequence).

Finally, it is important to realise that, although we presented the axioms for the underlying logic (the formulas φ_i that we assumed to be canonical) and the relational atoms as two independent layers, let us recall that [20] showed that interaction properties between the modalities and the relational atoms may be automatically 'imported'. For the case of epistemic logic $\mathbf{S5}$ with at least two agents and the $\text{Sup}(a, b)$ atom, one can derive that $\text{Sup}(a, b) \rightarrow K_b \text{Sup}(a, b)$ and $\neg \text{Sup}(a, b) \rightarrow K_b \neg \text{Sup}(a, b)$, in other words, in such a logic, it is derivable that if agent a considers at least the states possible that b considers possible, then b knows this! Similarly, if there is a state considered possible by b but not by a , then agent b knows this as well!

To summarise, we have presented a flexible way to deal locally with quantification over formulas. In particular, we have shown how, under some mild conditions, in a modal logic that extends \mathbf{K} with some canonical axioms, one can add a number of relational atoms, for each of them an axiom and an inference rule, such that the logic is complete for the class of models that interpret the atom as a first order property of the underlying frame. We argued that this presents many opportunities to express properties concerning the knowledge or beliefs of agents in a local way, so that they are only true now, or as a belief or knowledge of some specific agents. Although we focussed on epistemic and doxastic logics, our technique is applicable in temporal and dynamic settings as well. On our agenda is to study how our framework behaves in a dynamic epistemic logic setting. For instance, one might consider the effect of publicly announcing relational atoms, like $\text{Sup}(a, b)$, which would mean that it is announced that a knows at least what b knows. After such an announcement, one would expect that the local property

becomes global again, in many cases it would become a validity in the resulting model that $\forall y(R_b xy \Rightarrow R_a xy)$.

Like we explained, our completeness proof borrows ideas from both [16] and [5]. Also, the inference rule \mathbf{R}_{\Box} is reminiscent of an inference rule for irreflexivity [9]. However, it is important to stress that the approaches mentioned aim to axiomatise *global* properties. As far as we know, the work presented in this paper is a first general approach to *local* properties in models.

We thank the AAMAS reviewers for their helpful comments.

5. REFERENCES

- [1] A. Baltag, L. S. Moss, and S. Solecki. The logic of common knowledge, public announcements, and private suspicions. *TARK 98*, pp. 43–56, 1998.
- [2] P. Blackburn, J. van Benthem, and F. Wolter, (eds). *Handbook of Modal Logic*. Elsevier, Amsterdam, 2007.
- [3] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, volume 53. CUP, Cambridge, 2001.
- [4] P. Cohen and H. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.
- [5] M. de Rijke. *Extending Modal Logic*. PhD thesis, ILLC, University of Amsterdam, 1993.
- [6] K. Engelhardt, R. van der Meyden and Y. Moses. Knowledge and the logic of local propositions. *TARK 98*, pp. 29–41, 1998.
- [7] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning About Knowledge*. MIT Press, 1995.
- [8] M. Fitting. Modal proof theory. In Blackburn et al. [2], pp. 85 – 138.
- [9] D. Gabbay. An irreflexivity lemma with applications to axiomatisations of conditions on tense frames. In *Aspects of Philosophical Logic*, pp. 67–89. Reidel, 1981.
- [10] J. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *JACM*, 37(3):549–587, 1990.
- [11] D. Harel. Dynamic logic. In *Handbook of Philosophical Logic Vol. II*, pp. 497–604, 1984.
- [12] J. Hintikka. *Knowledge and Belief*, 1962.
- [13] J.-J. C. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. CUP, 1995.
- [14] R. C. Moore. Reasoning about knowledge and action. In *IJCAI-77*, Cambridge, MA, 1977.
- [15] A. S. Rao and M. P. Georgeff. Modeling rational agents within a BDI-architecture. In *KR&R-91*, pp. 473–484, 1991.
- [16] G. Renardel de Lavalette, B. Kooi, and R. Verbrugge. Strong completeness for PDL. In *AiML2002*, pp. 377–393, 2002.
- [17] J. van Benthem. *Modal Correspondence Theory*. PhD thesis, University of Amsterdam, 1976.
- [18] W. van der Hoek and M. Pauly. Modal logic for games and information. [2], pages 1077–1148.
- [19] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, Berlin, 2007.
- [20] H. van Ditmarsch, W. van der Hoek, and B. Kooi. Knowing more — towards a local correspondence theory. In *IJCAI-09*, pp. 955–960, 2009.
- [21] M. Wooldridge. *An Introduction to Multi-agent Systems*. John Wiley & Sons, 2002.